

## What's in the gap? Interaction transitions that make HRI work

Helge Hüttenrauch, Kerstin Severinson Eklundh, Anders Green, Elin A. Topp, Henrik I. Christensen

**Abstract** — This paper presents an in-depth analysis from a Human Robot Interaction (HRI) study on spatial positioning and interaction episode transitions. Subjects showed a living room to a robot to teach it new places and objects. This joint task was analyzed with respect to organizing strategies for interaction episodes. Noticing the importance of transitions between interaction episodes, small adaptive movements in posture were observed. This finding needs to be incorporated into HRI modules that plan and execute robots' spatial behavior in interaction, e.g., through dynamic adaptation of spatial formations and distances depending on interaction episode.

### I. INTRODUCTION

Socially interactive robots that operate in close proximity of humans and interact with them need to move in a predictable and appropriate way [1]. In interactive tasks that are performed in office and domestic environments the robot and its user initially start up their interaction, sustain it during the actual task, and finally end the mission. During such a joint interaction both participants need to carefully track each other's spatial configuration as part of the embodied interaction [2].

To guide the design of spatial robot behavior as part of the Cogniron project [3] we conducted an HRI study where users were instructed to make a robot follow through a "living room" like environment to show the robot new places and objects. Users were told that the robot had just been purchased and that it therefore had to go through this initial training and teaching to perform service missions later. To test whether the robot was familiar with these objects and locations the robot could be sent to find and verify these places and objects again.

The aim of the study was to explore and observe the posture and positioning patterns during a joint HRI task and to analyze observed user behaviors. The experiences gained from analyzing the interaction for spatial patterns and preferences are to be used to better understand and guide the design of robot behaviors in regard to approach-, motion-, and spatial adjustments to increase usability in HRI. The "home-tour" study was conducted at the Royal Institute of Technology with 22 subjects. The *static spatial relationships* re-



Fig.1. User closely observing the robot during the HRI trial

garding the preferred distance and formation of users in relation to the robot according to Hall's interpersonal distance zones and Kendon's F-formation system are reported in [4]<sup>1</sup>.

However, we also discovered during this first analysis that not only the stationary parameters of posture and positioning were of importance, but that the *dynamic positioning changes* seemed to play an important role in signaling the change from one interaction episode to the next. To study this user behavior in detail we analyzed our available trial data focusing on the transitions between *interaction episodes*. The questions we asked were:

- How do users organize the task of showing the robot its new environment?
- Are preferences of user behavior recognizable that might aid the design of spatial behaviors for HRI?
- Can interaction episode transition patterns be detected automatically and used for spatial management in a joint HRI task?

Answers to these questions will make it possible to design improved robot control systems that take into account the social and communicative aspects of managing space based upon user behavior and users' expectations.

The remaining paper is organized as follows. The background to managing space in HRI and related research in robotics is given in the following section. In Section III the user study conducted is summarized briefly before the

Manuscript submitted March 21, 2006. The work described in this paper was conducted within the EU Integrated Project COGNIRON ('The Cognitive Robot Companion', [www.cogniron.org](http://www.cogniron.org)) and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020. 002020.

H. Hüttenrauch, K. Severinson Eklundh, A. Green, E. A. Topp, and H. I. Christensen are with the Royal Institute of Technology (KTH), School for Computer Science and Communication (CSC), SE-100 44, Stockholm, Sweden; (corresponding author: H. Hüttenrauch: +46 8-790 9156; fax: +46 8-790 90 99; e-mail: [hehu@csc.kth.se](mailto:hehu@csc.kth.se)).

<sup>1</sup> available as "IPLab-260", technical report, see: <http://ftp.nada.kth.se/IPLab/TechReports/IPLab-260.pdf>

analysis on dynamic transitions is presented and described. Finally, in section IV we discuss our findings.

## II. MANAGING SPACE IN HRI

Multiple perspectives of managing the shared space in human robot interaction have been investigated. One aspect is the effort to design for safe robot behaviors and interactions [5], [6], [7]. Possible sensor systems enabling a safe interaction are studied as well as algorithms to detect and recognize human movements [8], [9]. Other publications are focusing on navigation and localization (planning) in the presence of humans [10].

The tested human robot interaction situations, robotic strategies, and systems vary considerably and include for example a crowded subway station traversed with a robotic wheelchair and an accompanying person [11] or users passing a moving robot in a hallway [12]. A robot entering a room and joining a group of people was tried in [13]. Also reported was an autonomous robot that gets into a queue of people and advances in it to a service point [14].

Other studies focus on the *user behavior and preferences* in human robot interaction: Butler and Agah [15] varied a robot's movement behaviors and performed a user study to evaluate how different robot speeds and distances were perceived by users. More recently, Walters et al. [16] tried to relate preferences in social distance for humans approaching a robot and – vice versa – robots approaching humans with their subjects' personality traits and found a correlation between the social distance and a factor termed "Proactiveness" by the authors. In another HRI experimental study [17] the preferred approach strategy of a robot coming towards a *sitting user* was investigated. A robot brought a remote control and a majority of subjects preferred to have a robot approach from the *right* side (from the users' point of view); least liked was the robot approaching directly from a frontal positioning (or in vis-à-vis formation).

In [4] findings from human to human interaction studies, i.e., Hall's *interpersonal distances* [18] and Kendon's *F-formation system arrangements* [19] were investigated in a human robot interaction scenario. Quantitatively describing the interaction-initiating distances and positioning, the authors reported a preference for the *vis-à-vis positioning* and a preferred distance in the range of Hall's *personal distance* (i.e., 0.46 – 1.22 meters) while initiating a robot mission. However, it was also cautioned that these findings are characterized by their *static parameter quality*, i.e., treating and describing the interaction between a user and a robot as a singular point-in-time only, ignoring the *dynamic flow* of interaction altogether.

## III. EXPERIMENTAL HRI STUDY AND FINDINGS

The HRI study analyzed for spatial posture and positioning changes during transitions between interaction episodes has previously been reported in detail, e.g., with regard to interpersonal distances and F-formations [4] and miscom-

munication in the spoken dialogue [20]. Consequently the study's set-up will here only be summarized briefly. Instead the method of analyzing and the presentation of findings on the dynamic transitions between interaction episodes will be focused upon.

### A. HRI study: Set-up and data collection

We investigated the multimodal interaction and spatial positioning in a HRI usage scenario termed the "Home Tour". This scenario explores the joint human-robot task of showing a robot around in an environment and teaching it places and objects by naming them. Additionally, subjects could validate previously taught items by sending the robot to find them again. The study was conducted in a laboratory that is furnished like a living room. For the trial 22 subjects from the Royal Institute of Technology were recruited to participate. A commercially available PeopleBot<sup>2</sup> was used in the study. Its movement and its on-board-camera as well as the generated speech-dialogue output were remotely controlled according to the Wizard-of-Oz methodology [21].

Data collected included the video from an external video camera, audio recording on the robot, four webcams in the room's corners to take still images, data from the robot's laser range finder (LRF), a short user questionnaire as well as the systems' files of logged commands and system states.

This rich set of data has been transcribed, processed, and analyzed with different research objectives in a frame-by-frame analysis for 11 trials so far. The data were synchronized so that the different media could be analyzed together. The speech dialogue was annotated so that the mission task progress, observed gestures, postures, and positioning changes could be described and commented. This annotation required the analysis of the different AV-recordings, still images, and data sets, e.g., from the LRF. The analysis resulted in a corpus of HRI instances that is available in two different media types: One is a text based, printable document format (that can be machine-read however). The other media type is a computer based interactive multimedia representation produced with Anvil [22]. Anvil allows representing different tracks of analysis in a music-score like, parallel representation that can be annotated and played back with digitized video and other digital media formats.

### B. Transition – the dynamic event

Investigating the initial, but static spatial positioning and interpersonal distances between trial subjects and the robot we noticed that a helpful differentiation of the HRI units of analysis can be found in what we termed *interaction episodes*. An interaction episode in a joint HRI task has a clearly marked start – often by speech dialogue utterance (e.g., an issued command) – or observable actions taken by either the robot or its (trial) user signifying such a beginning. An ongoing interaction episode is delimited by the start of the following interaction episode. In our trial sce-

<sup>2</sup> [www.mobilerobots.com/](http://www.mobilerobots.com/)



nario the interaction episodes of “FOLLOW” (user guiding the robot around), “SHOW” (user teaching the robot places and objects), and “VALIDATE” (user testing the taught places and objects by sending the robot to them again) were identified and used for annotation and analysis.

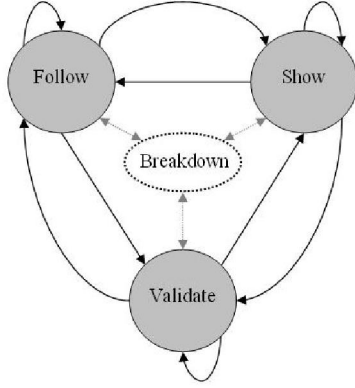


Fig. 2. Interaction Episodes and possible Transitions

The relationship of the interaction episodes in our “Home tour” HRI scenario can thus be illustrated as shown in fig. 2: Each of the three states can be followed by a *transition* into an interaction episode. Furthermore, in case of a communicative or interaction “failure”, an additional state termed BREAKDOWN (including a possible repair) can be entered. An example of such a transition from a Follow into the Breakdown state could be, e.g., a trial user asking the robot to follow and then too quickly moving around the robot so that the robot can not “see” the user any longer. When such an incident was observed, the robot would announce that Following has stopped and request the user to come back in front of the robot.

The arrows connecting the different states are of importance for HRI analyses and understanding as they symbolize the *dynamic changes* or *transitions* occurring when a user and a robot want to enter into a different episode during interaction. Both interacting partners need to switch their attention to this “new topic” and get in a joint spatial configuration that best suits or is preferred for this next phase.

### C. Observed transitions in the study

We analyzed 11 transcribed trials in detail to quantify the transitions observed and to study the joint spatial behaviors of the user and the robot observed in the different data sources from the trial.

The interaction episodes Follow, Show, and Validate were identified in the data according to the above described delimitation of speech utterances, body movement like gesture and posture changes, and actions taken in the environment. Excluded from this discussion are BREAKDOWN states and transitions as their understanding and just treatment are beyond the scope of this article – details on the communicative aspects of miscommunication are discussed further in [20].

The descriptive analysis of the interaction episode transi-

tions is summarized in table 1 for a total of  $N=294$  transitions found in 11 HRI trials. Most transitions happened between the Follow and the Show state ( $n_{FS} = 79$ ), and in the opposite direction again, i.e., from Show to Follow with  $n_{SF} = 53$ . More Show than Follow states lead to a Validate interaction episode ( $n_{SV} = 35$ , compared to  $n_{FV} = 10$ ). Furthermore it is interesting to look at the “more-of-the-same” transitions, i.e., an interaction episode that is followed by another interaction episode of the same kind (depicted in underlined style in table 1). While very few Follow missions are leading to another Follow mission ( $n_{FF} = 7$ ), a second Validate mission is comparably more often initiated after a previous one ( $n_{VV} = 48$ ).

TABLE I: NUMBER OF INTERACTION EPISODE TRANSITIONS, ( $N=294$ )

Transition From:	To:		
	Follow	Show	Validate
Follow	<u>7</u>	79	10
Show	53	<u>26</u>	35
Validate	27	9	<u>48</u>

In summary these findings suggest that there exists a logical order in the chain of different interaction episodes and transitions between them based on the experimental task and instruction given to subjects: Follow and Show states are first iterated to explain to the robot different places and objects. A Validate mission in turn is more likely to be started after a finished Show than a Follow mission. The transition from the Validate state displays two interesting observations: Validate states are often followed by another Validate interaction. Additionally, once the Validate mission state is left, the “natural” order is to start over with another Follow subtask. Going back to the videos of the trials we found two possible reasons for the observations regarding the Validate transitions into other states. Subjects sometimes tried repeatedly to test the robot with different Validate missions that the robot was unable to fulfill, e.g., finding objects that the user had previously taken and placed somewhere else. When this “searching” for such an object failed repeatedly, subjects sometimes simply abandoned the Validate mission and re-started the interaction with the robot by issuing a new Follow mission.

### D. Small spatial moves as transition cues

During our trial annotation and analysis we noticed that there seemed to be small movements just before or during the transition to another interaction episode. To study this finding of small spatial adjustment movements more closely we looked especially for this behavior, e.g., in the robot’s laser-range finder sensory data as well as the audiovisual sources of trial recordings.

One example of such a transition and adjustment movements is given in detail in figure 3, 4 and the image series (figure 5-8) below. However, as we found this behavior on multiple transition occasions we see the given example as a

prototype of the small spatial adaptation found during transitions between interaction episodes.

Figure 3 and 4 are both illustrations of the laser range finder data. The recorded sensory input was processed and analyzed by a human detection and tracking system [9] to reveal the robot-centric view of the relative distance and angle between the robot and its subject. Figure 3 shows how the distance changes while the subject makes a transition from a Follow state into a Show state of interaction. Figure 4 depicts the same time-interval for the relative angle between the robot and the subject.

The figures 5 to 8 show four “time-slices” of this transition. Numbered ❶ to ❸ below, they will be used in illustrating the step-wise explanation of the transition and the small spatial adjustments observed.

The subject stands near a table and awaits the approaching robot as part of a Follow interaction episode. The approach can clearly be seen from the decreasing distance (fig.3) and the slightly changing angle (fig.4) between the robot and the subject. The user stops the robot by commanding “Stop robot”. This spoken command from the subject and the robot’s acknowledging it (“Stopped following”) ends the previous Follow episode as discussed above.

Note that the next episode is not immediately started, i.e., the “Show-2” episode starts a couple of seconds after the previous episode has ended (depicted in fig. 3 and 4 by the non-connecting classification-boxes). The cause for this “delay” can be seen in illustration❶: The user hesitates what to do next and checks the trial’s instruction sheet once more. Then the subject decides to show and teach the table as object to the robot. The subject *prepares* for this Show episode by stepping towards the table *shortly before* taking up the pointing gesture and saying “This is a table.” as seen in illustration❷. The start of the pointing gesture and the speech utterance is taken as indicator that the next interaction episode (“Show-2”) has started.

Interesting to the question of spatial management behavior is the subject’s small movement and change in distance and angle towards the robot that allows the subject to point towards the table: The distance for example changes from 50.2 cm to 64.5 cm as pointed out by the vertical arrow in figure 3. This is a seemingly small distance change, however, we find this *small adaptation movement as preparatory move* interesting as it (1) can be taken as an indication that a transition towards another interaction episode is occurring, and (2) can be “seen” by the robot’s sensory system, i.e., a robot can detect this change in spatial positioning.

The aspect and importance of the robot’s spatial management can also be discussed. The robot came to a stop next to the subject as consequence of the previous Follow mission, i.e., relative to the subject’s previous positioning when the Follow mission was ordered to be terminated. The subject then made a move to show the table. However, the robot does not follow autonomously in positioning or orientation with the user. Consequently it is difficult for the robot to

truly “see” the table that is now being pointed to by the subject.

The navigating (remote controlling) “Wizard” of the robot detects this shortcoming in the robot’s positioning and changes without request from the subject the orientation of the robot (see horizontal arrow in figure 4 indicating the changing angle towards the user, or the robot’s rotating towards the subject in illustration image❸). This action from the experiment leader is due to the *spatial adaptation* that the robot should perform according to the technical requirement of orientating the robot towards a user and the shown object in a Show episode (as seen in illustration❹).

#### IV. DISCUSSION

We have reported a HRI study targeted at a “Home Tour” scenario to teach a robot its new operation area and objects within it. The analysis of the interaction observed introduced a categorization with help of different *interaction episodes*, i.e., Follow-, Show-, and Validate-states. They can be used as multimodal HRI units delimiting and structuring the flow of actions. Interaction episodes can be identified by looking at the tasks performed, the conversation in spoken language as well as the posture and positioning changes during an interaction between a robot and a user.

Inspired by previously reported findings from this study the dynamic aspects of *transitions* between different interaction episodes was explained and quantified. Users organize the task of showing the robot around, and patterns of user behavior can be detected by looking at the distribution of transitions between different interaction episodes. Examples for such identified patterns are the iteration of Follow and Show states or the multiple initiations of Validate missions after one another.

The relevance of spatial management in transitions has been illustrated by showing how a subject may *prepare for a following* interaction episode by slightly adapting his position and orientation towards the robot or the environment *shortly before or during such an interaction episode transition*. This preparing movement can also be interpreted as a sign of *co-operation* from the user [23]. By signaling the intent of the coming action with a small movement adjustment a willingness to make this next action understandable for the interaction partner can be anticipated. We feel that a more detailed investigation of this co-operative behavior is called for.

As the robot’s LRF can be used to detect such small adjustment changes signifying the switch from one interaction episode to another it seems like a promising approach to try to incorporate such detection and necessary spatial movement and orientation adjustment behaviors into the robots’ navigation system and interaction planner components. Our findings on the importance of transitions of interaction episodes for the spatial management can be seen as an instance of what Schegloff [24] observed in human to human interaction and termed “body torque”, i.e., how the

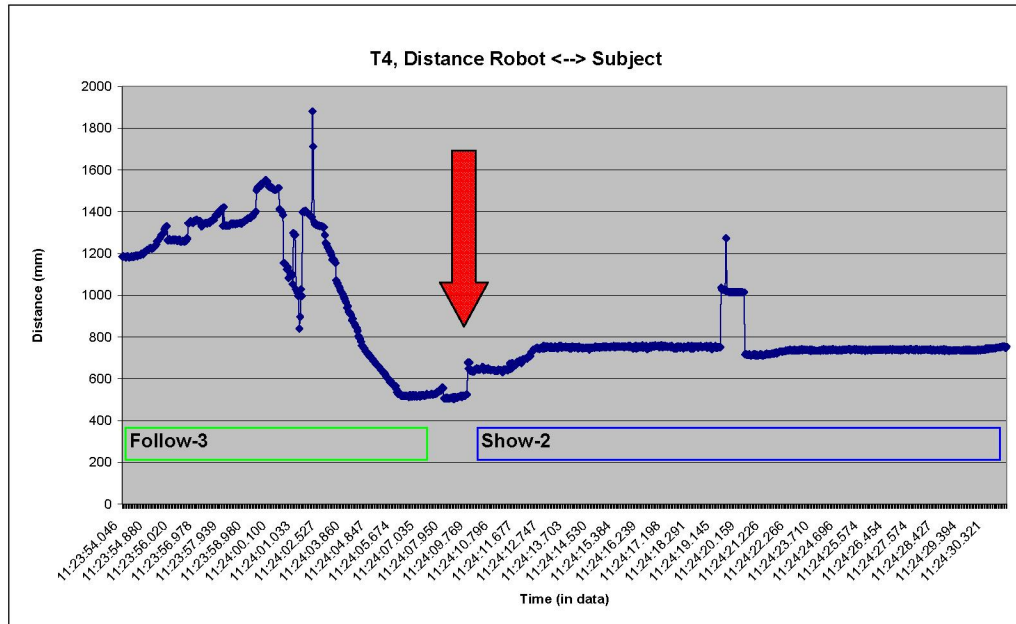


Fig. 3. Trial 4; Distance between Robot and Subject – at arrow, transition from Follow to Show episode

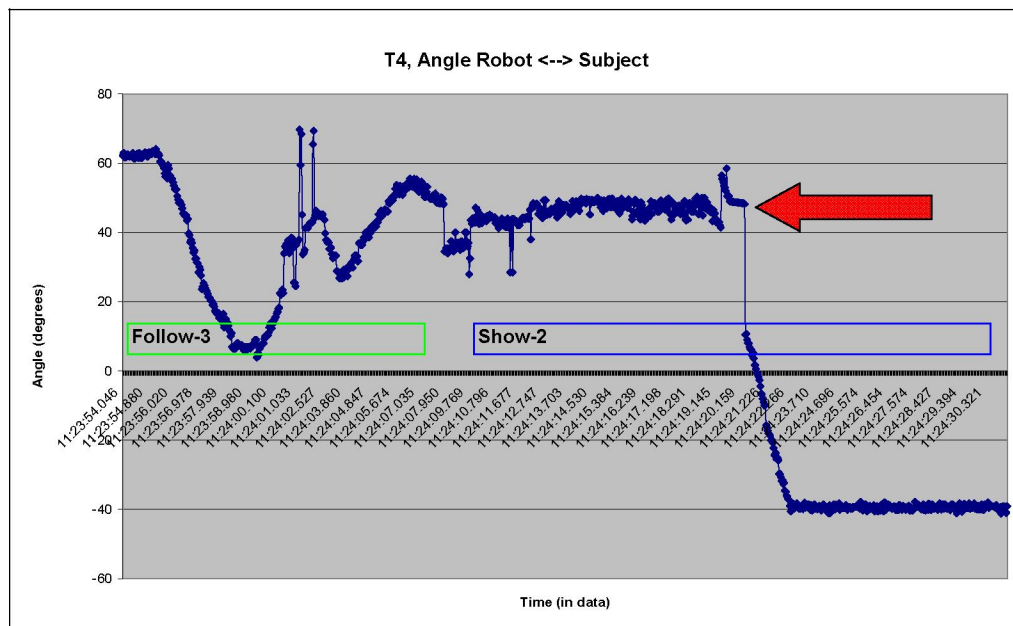


Fig. 4. Trial 4, Angle between Robot and Subject – at arrow, robot's orientation change towards user



Fig. 5-8. (left to right): ① Follow stopped; ② User in new position; ③ Robot rotates towards user; ④ Show episode continued

direction and orientation of, e.g., the upper body in relationship to the lower extremities are able to shape an ongoing interaction as well as give a potential outlook how this interaction is expected to develop in the next phase by its participants.

We have recently started to make use of this effect by trying an improved robot signaling model in transitions between interaction episodes. Termed “*spatial prompting*” [25] we would like to test robot behaviors that exhibit such small, adjusting robot movements to test whether they are sufficient to better support users’ understanding of the current state in interaction. Results of a head-nodding robot giving gestural feedback are encouraging support to this idea [26]. If spatial prompting behaviors can be demonstrated to work, we expect that the number of communicative breakdowns or miscommunication can be decreased, leading to an improved robot usage experience.

In judging the validity of our research method and reported findings it is important to consider that the robot in speech interaction, robot navigation, and camera movement of the on-board camera was controlled by a human operator (or “Wizard” according to [21]). Consequently the observed movements from the robot are not necessarily the ones that really implemented robot navigation and interaction components might exhibit in interacting with users or reacting to users’ movement and communicative behavior. Furthermore our study was confined to a single room of about 5 x 5 meters in size, furnished “living room”-like, but not in all resembling a living room at home.

As consequence of these limitations we have recently started to change a few attributes of our study methodology: We have exchanged previously used remote control systems, e.g., for navigation, with an autonomous robot navigation and localization, and for the following of users a working human detection and tracking system has been implemented. Additionally we have extended the operation area from a single room to a multiple-room scenario which gives additional interesting situations and possible interaction episode transitions, e.g., a user and a robot passing a narrow passage together.

## REFERENCES

- [1] T. Fong, I. Nourbakhsh and K. Dautenhahn, “A survey of socially interactive robots”. *Robotics and Autonomous Systems*, 42, 2003; pp. 143-166.
- [2] P. Dourish. *Where The Action Is: The Foundations of Embodied Interaction*. MIT Press, 2001.
- [3] R. Chatila, “The Cognitive Robot Companion and the European ‘Beyond Robotics Initiative’”, in: *6th EAJ International Symposium “Living with Robots”*, 2004.
- [4] H. Hüttenrauch, K. Severinson Eklundh, A. Green, and E. A. Topp. “Investigating Spatial Relationships in Human-Robot interaction”. To be published in: *Proc. of the 2006 IEEE/RSJ Intern. Conference on Intelligent Robots and Systems (IROS 2006)*.
- [5] M. Yoda, Y. Shiota. “The mobile robot which passes a man”. In: *Proc. of the IEEE Intern. Workshop on Robot and Human Interactive Communication (ROMAN)*, 1997, pp. 112-117.
- [6] S. Nonaka, K. Inoue, T. Arai, and Y. Mae. “Evaluation of human sense of security for coexisting robots using virtual reality 1 st report: evaluation of pick and place motion of humanoid robots.” In: *Proc. of Robotics and Automation (ICRA ’04)*. 2004 IEEE Intern. Conference on, Volume:3, 2004. pp.2770-2775.
- [7] V.J. Traver A.P. del Pobil, M. Perez-Francisco. “Making service robots human-safe”. In: *Intelligent Robots and Systems (IROS 2000)*, Proceedings 2000 IEEE/RSJ International Conference on, Vol.1, pp. 696-701.
- [8] M. Finke et al. “Hey, I’m over here - How can a robot attract people’s attention?.” in: *Robot and Human Interactive Communication (ROMAN 2005)*, IEEE International Workshop on, 2005, pp. 7-12.
- [9] E.A. Topp and H.I. Christensen. “Tracking for Following and Passing Persons”. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2005)*, 2005.
- [10] K. Madhava Krishna, R. Alami, T. Simeon. “Safe proactive plans and their execution”. *Robotics and Automation System*, to be published.
- [11] E. Prassler, D. Bank, B. Kluge. “Motion Coordination between a Human and a Mobile Robot.” in: *Proc. of the 2002 IEEE/RSJ Intern. Conference on Intelligent Robots and Systems (IROS’02)*, Lausanne, Switzerland, 2002.
- [12] E. Pacchierotti, H.I. Christensen, P. Jensfelt. “Human-Robot Embodied Interaction in Hallway Settings: A Pilot study.” in: *IEEE Intern. Workshop on Robots and Human Interactive Communication (ROMAN)*, 2005, pp. 164-171.
- [13] P. Althaus, H. Ishiguro, T. Kanda, T. Miyashita, H.I. Christensen. “Navigation for human-robot interaction tasks”. In: *Proc. of the IEEE International Conference on Robotics and Automation (ICRA’04)*, 2004, pp. 1894-1900.
- [14] Y. Nakauchi, R. Simmons. “A social robot that stands in line” in: *Proc. of the IEEE/RSJ International Conference on Intelligent Robot and Systems (IROS)*, 2000, pp. 357-364.
- [15] J.T. Butler, A. Agah. “Psychological effects of behavior patterns of a mobile personal robot”. *Autonomous Robots*, Vol.10, pp. 185-202, March 2001.
- [16] M.L. Walters et al. “The influence of subjects’ personality traits on personal spatial zones in a human-robot interaction experiment”. in *Proc. 14th IEEE Int. Workshop on Robot & Human Communication (RO-MAN)*, 2005, pp. 347-352.
- [17] K. Dautenhahn et al. “How May I Serve You? A Robot Companion Approaching a Seated Person in a Helping Context.” In: *Proceedings of the Human Robot Interaction Conference (HRI’06)*, ACM Press, 2006. pp. 172-179.
- [18] E. T. Hall. *The Hidden Dimension: Man’s Use of Space in Public and Private*. The Bodley Head Ltd, London, UK, 1966.
- [19] A. Kendon. *Conducting interaction – Patterns of behavior in focused encounters*. Studies in interactional sociolinguistics. Cambridge, 1990.
- [20] A. Green, B. Wrede, K. Severinson Eklundh, and S. Li.. “Integrating Miscommunication Analysis in the Natural Language Interface Design for a Service Robot”. To be published in: *Proc. of the 2006 IEEE/RSJ Intern. Conference on Intelligent Robots and Systems (IROS 2006)*.
- [21] A. Green, H. Hüttenrauch, K. Severinson-Eklundh. “Applying the Wizard-of-Oz Framework to Cooperative Service Discovery and Configuration”. in *Proc. of the 13th IEEE International Workshop on Robot and Human Interactive Communication*, 2004.
- [22] M. Kipp. *Gesture Generation by Imitation - From Human Behavior to Computer Character Animation*. Dissertation.com, Boca Raton, Florida, 2004.
- [23] S.P. Gill, M. Kawamori, Y. Katagiri, A. Shimojima. “The Role of Body Moves in Dialogue.” In *International Journal of Language and Communication (RASK)*, Vol.12 (April), 2000.
- [24] E. Schegloff, “Body Torque.” *Social Research* 65, 3 (1998): 535-596.
- [25] A. Green, H. Hüttenrauch, “Making a Case for Spatial Prompting in Human-Robot Communication.”, Multimodal Corpora: From Multimodal Behaviour theories to usable models, *Workshop at the 5th Intern. Conf. on Language Resources and Evaluation (LREC 2006)*.
- [26] C. L. Sidner, C. Lee, L-P. Morency, C. Forlines. “The Effect of Head-Nod Recognition in Human-Robot Conversation”. In: *Proc of HRI’06*, 2006. ACM Press, New York, NY. pp: 290-296.